

RECEIVED  
JUL 10 2003  
TECH CENTER 1600



1600

## RAW SEQUENCE LISTING

DATE: 07/07/2003

PATENT APPLICATION: US/09/051,034B

TIME: 14:36:20

Input Set : A:\305626u1.app

Output Set: N:\CRF4\07072003\I051034B.raw

3 <110> APPLICANT: MCKENZIE, IAN FARQUHAR CAMPBELL  
4 SANDRIN, MAURO SERGIO  
6 <120> TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC  
7 GLYCOSYLTRANSFERASE  
9 <130> FILE REFERENCE: 30562.6USWO  
11 <140> CURRENT APPLICATION NUMBER: 09/051,034B  
12 <141> CURRENT FILING DATE: 1998-03-31  
14 <150> PRIOR APPLICATION NUMBER: PCT/AU97/00492  
15 <151> PRIOR FILING DATE: 1997-08-01  
17 <150> PRIOR APPLICATION NUMBER: 60/024,279  
18 <151> PRIOR FILING DATE: 1996-08-21  
20 <150> PRIOR APPLICATION NUMBER: PO1402  
21 <151> PRIOR FILING DATE: 1996-08-02  
23 <160> NUMBER OF SEQ ID NOS: 16  
25 <170> SOFTWARE: PatentIn Ver. 2.1  
27 <210> SEQ ID NO: 1  
28 <211> LENGTH: 1043  
29 <212> TYPE: DNA  
30 <213> ORGANISM: Sus Domesticus  
32 <220> FEATURE:  
33 <221> NAME/KEY: CDS  
34 <222> LOCATION: (9)..(1031)  
36 <400> SEQUENCE: 1

37 ctacagcc atg ctc agc atg cag gca tcc ttc ttc ttc ccc acg ggt ccc 50  
38 Met Leu Ser Met Gln Ala Ser Phe Phe Phe Pro Thr Gly Pro  
39 1 5 10

41 ttc atc ctc ttt gtc ttc acg gct tcc acc ata ttt cac ctt cag cag 98  
42 Phe Ile Leu Phe Val Phe Thr Ala Ser Thr Ile Phe His Leu Gln Gln  
43 15 20 25 30

45 agg atg gtg aag att caa ccc acg tgg gag tta cag atg gtg acg cag 146  
46 Arg Met Val Lys Ile Gln Pro Thr Trp Glu Leu Gln Met Val Thr Gln  
47 35 40 45

49 gtg acc aca gag agc ccc tcg agc ccc cag ctg aag ggc atg tgg acg 194  
50 Val Thr Thr Glu Ser Pro Ser Ser Pro Gln Leu Lys Gly Met Trp Thr  
51 50 55 60

53 atc aat gcc atc ggc cgc ctg ggg aac cag atg ggg gag tac gcc acc 242  
54 Ile Asn Ala Ile Gly Arg Leu Gly Asn Gln Met Gly Glu Tyr Ala Thr  
55 65 70 75

57 ctg tac gcg ctg gcc agg atg aac ggg cgg ccg gcc ttc atc ccg ccc 290  
58 Leu Tyr Ala Leu Ala Arg Met Asn Gly Arg Pro Ala Phe Ile Pro Pro  
59 80 85 90

61 gag atg cac agc acg ctg gcc ccc atc ttc agg atc acc ctc ccg gtc 338  
62 Glu Met His Ser Thr Leu Ala Pro Ile Phe Arg Ile Thr Leu Pro Val

p.6

ENTERED

## RAW SEQUENCE LISTING

DATE: 07/07/2003

PATENT APPLICATION: US/09/051,034B

TIME: 14:36:20

Input Set : A:\305626u1.app

Output Set: N:\CRF4\07072003\I051034B.raw

```

63 95          100          105          110
65 ctg cac gcc agc acg gcc cgc agg atc ccc tgg cag aac tac cac ctg 386
66 Leu His Ala Ser Thr Ala Arg Arg Ile Pro Trp Gln Asn Tyr His Leu
67          115          120          125
69 aac gac tgg atg gag gag cgg tac cgc cac atc ccg ggg gag tac gtg 434
70 Asn Asp Trp Met Glu Glu Arg Tyr Arg His Ile Pro Gly Glu Tyr Val
71          130          135          140
73 cgc ctc acg ggc tac ccc tgc tcc tgg acc ttc tac cac cac ctg cgc 482
74 Arg Leu Thr Gly Tyr Pro Cys Ser Trp Thr Phe Tyr His His Leu Arg
75          145          150          155
77 acc gag atc ctc cgg gag ttc acc ctg cat aac cac gtg cgc gag gag 530
78 Thr Glu Ile Leu Arg Glu Phe Thr Leu His Asn His Val Arg Glu Glu
79          160          165          170
81 gcc cag gat ttc ctg cgg ggt ctg cgg gtg aac ggg agc cga ccg agt 578
82 Ala Gln Asp Phe Leu Arg Gly Leu Arg Val Asn Gly Ser Arg Pro Ser
83 175          180          185          190
85 acc tac gtg ggg gtg cac gtg cgc cgg ggg gac tac gtg cac gtg atg 626
86 Thr Tyr Val Gly Val His Val Arg Arg Gly Asp Tyr Val His Val Met
87          195          200          205
89 ccc aac gtg tgg aag ggc gtg gtg gcc gac cgg cgg tac ctg gag cag 674
90 Pro Asn Val Trp Lys Gly Val Val Ala Asp Arg Arg Tyr Leu Glu Gln
91          210          215          220
93 gcc ctg gac tgg ttc cgg gct cgc tac cgc tcc ccc gtc ttt gtg gtc 722
94 Ala Leu Asp Trp Phe Arg Ala Arg Tyr Arg Ser Pro Val Phe Val Val
95          225          230          235
97 tcc agc aac ggc atg gcc tgg tgt cgg gaa aac atc aat gcc tcg cgc 770
98 Ser Ser Asn Gly Met Ala Trp Cys Arg Glu Asn Ile Asn Ala Ser Arg
99          240          245          250
101 ggc gat gtg gtg ttt gcc ggc aat ggc atc gag ggc tcc ccc gcc aaa 818
102 Gly Asp Val Val Phe Ala Gly Asn Gly Ile Glu Gly Ser Pro Ala Lys
103 255          260          265          270
105 gac ttc gcg ctg ctc acg cag tgt aac cac act gtc atg acc att ggc 866
106 Asp Phe Ala Leu Leu Thr Gln Cys Asn His Thr Val Met Thr Ile Gly
107          275          280          285
109 acg ttc ggg atc tgg gcc gcc tac ctt gct ggt gga gag acc atc tac 914
110 Thr Phe Gly Ile Trp Ala Ala Tyr Leu Ala Gly Gly Glu Thr Ile Tyr
111          290          295          300
113 ctg gcc aat tac acg ctc ccg gac tct ccc ttc ctc aaa ctc ttt aag 962
114 Leu Ala Asn Tyr Thr Leu Pro Asp Ser Pro Phe Leu Lys Leu Phe Lys
115          305          310          315
117 ccc gag gca gcc ttc ctg ccc gag tgg att ggg atc gag gca gac ctg 1010
118 Pro Glu Ala Ala Phe Leu Pro Glu Trp Ile Gly Ile Glu Ala Asp Leu
119          320          325          330
121 tcc cca ctc ctt aag cac tga tgcggtgt cc 1043
122 Ser Pro Leu Leu Lys His
123 335          340
126 <210> SEQ ID NO: 2
127 <211> LENGTH: 340
128 <212> TYPE: PRT

```

## RAW SEQUENCE LISTING

DATE: 07/07/2003

PATENT APPLICATION: US/09/051,034B

TIME: 14:36:20

Input Set : A:\305626u1.app

Output Set: N:\CRF4\07072003\I051034B.raw

129 &lt;213&gt; ORGANISM: Sus Domesticus

131 &lt;400&gt; SEQUENCE: 2

```

132 Met Leu Ser Met Gln Ala Ser Phe Phe Phe Pro Thr Gly Pro Phe Ile
133   1           5           10           15
134 Leu Phe Val Phe Thr Ala Ser Thr Ile Phe His Leu Gln Gln Arg Met
135           20           25           30
136 Val Lys Ile Gln Pro Thr Trp Glu Leu Gln Met Val Thr Gln Val Thr
137           35           40           45
138 Thr Glu Ser Pro Ser Ser Pro Gln Leu Lys Gly Met Trp Thr Ile Asn
139           50           55           60
140 Ala Ile Gly Arg Leu Gly Asn Gln Met Gly Glu Tyr Ala Thr Leu Tyr
141   65           70           75           80
142 Ala Leu Ala Arg Met Asn Gly Arg Pro Ala Phe Ile Pro Pro Glu Met
143           85           90           95
144 His Ser Thr Leu Ala Pro Ile Phe Arg Ile Thr Leu Pro Val Leu His
145           100          105          110
146 Ala Ser Thr Ala Arg Arg Ile Pro Trp Gln Asn Tyr His Leu Asn Asp
147           115          120          125
148 Trp Met Glu Glu Arg Tyr Arg His Ile Pro Gly Glu Tyr Val Arg Leu
149           130          135          140
150 Thr Gly Tyr Pro Cys Ser Trp Thr Phe Tyr His His Leu Arg Thr Glu
151  145          150          155          160
152 Ile Leu Arg Glu Phe Thr Leu His Asn His Val Arg Glu Glu Ala Gln
153           165          170          175
154 Asp Phe Leu Arg Gly Leu Arg Val Asn Gly Ser Arg Pro Ser Thr Tyr
155           180          185          190
156 Val Gly Val His Val Arg Arg Gly Asp Tyr Val His Val Met Pro Asn
157           195          200          205
158 Val Trp Lys Gly Val Val Ala Asp Arg Arg Tyr Leu Glu Gln Ala Leu
159           210          215          220
160 Asp Trp Phe Arg Ala Arg Tyr Arg Ser Pro Val Phe Val Val Ser Ser
161  225          230          235          240
162 Asn Gly Met Ala Trp Cys Arg Glu Asn Ile Asn Ala Ser Arg Gly Asp
163           245          250          255
164 Val Val Phe Ala Gly Asn Gly Ile Glu Gly Ser Pro Ala Lys Asp Phe
165           260          265          270
166 Ala Leu Leu Thr Gln Cys Asn His Thr Val Met Thr Ile Gly Thr Phe
167           275          280          285
168 Gly Ile Trp Ala Ala Tyr Leu Ala Gly Gly Glu Thr Ile Tyr Leu Ala
169           290          295          300
170 Asn Tyr Thr Leu Pro Asp Ser Pro Phe Leu Lys Leu Phe Lys Pro Glu
171  305          310          315          320
172 Ala Ala Phe Leu Pro Glu Trp Ile Gly Ile Glu Ala Asp Leu Ser Pro
173           325          330          335
174 Leu Leu Lys His
175           340
179 <210> SEQ ID NO: 3
180 <211> LENGTH: 1098
181 <212> TYPE: DNA

```

## RAW SEQUENCE LISTING

DATE: 07/07/2003

PATENT APPLICATION: US/09/051,034B

TIME: 14:36:20

Input Set : A:\305626u1.app

Output Set: N:\CRF4\07072003\I051034B.raw

```

182 <213> ORGANISM: Sus Domesticus
184 <220> FEATURE:
185 <221> NAME/KEY: CDS
186 <222> LOCATION: (1)..(1098)
188 <400> SEQUENCE: 3
189 atg tgg gtc ccc agc cgc cgc cac ctc tgt ctg acc ttc ctg cta gtc 48
190 Met Trp Val Pro Ser Arg Arg His Leu Cys Leu Thr Phe Leu Leu Val
191 1 5 10 15
193 tgt gtt tta gca gca att ttc ttc ctg aac gtc tat caa gac ctc ttt 96
194 Cys Val Leu Ala Ala Ile Phe Phe Leu Asn Val Tyr Gln Asp Leu Phe
195 20 25 30
197 tac agt ggc tta gac ctg ctg gcc ctg tgt cca gac cat aac gtg gta 144
198 Tyr Ser Gly Leu Asp Leu Leu Ala Leu Cys Pro Asp His Asn Val Val
199 35 40 45
201 tca tct ccc gtg gcc ata ttc tgc ctg gcg ggc acg ccg gta cac ccc 192
202 Ser Ser Pro Val Ala Ile Phe Cys Leu Ala Gly Thr Pro Val His Pro
203 50 55 60
205 aac gcc tcc gat tcc tgt ccc aag cat cct gcc tcc ttt tcc ggg acc 240
206 Asn Ala Ser Asp Ser Cys Pro Lys His Pro Ala Ser Phe Ser Gly Thr
207 65 70 75 80
209 tgg act att tac ccg gat ggc cgg ttt ggg aac cag atg gga cag tat 288
210 Trp Thr Ile Tyr Pro Asp Gly Arg Phe Gly Asn Gln Met Gly Gln Tyr
211 85 90 95
213 gcc acg ctg ctg gcc ctg gcg cag ctc aac ggc cgc cag gcc ttc atc 336
214 Ala Thr Leu Leu Ala Leu Ala Gln Leu Asn Gly Arg Gln Ala Phe Ile
215 100 105 110
217 cag cct gcc atg cac gcc gtc ctg gcc ccc gtg ttc cgc atc acg ctg 384
218 Gln Pro Ala Met His Ala Val Leu Ala Pro Val Phe Arg Ile Thr Leu
219 115 120 125
221 cct gtc ctg gcg ccc gag gta gac agg cac gct cct tgg cgg gag ctg 432
222 Pro Val Leu Ala Pro Glu Val Asp Arg His Ala Pro Trp Arg Glu Leu
223 130 135 140
225 gag ctt cac gac tgg atg tcc gag gat tat gcc cac tta aag gag ccc 480
226 Glu Leu His Asp Trp Met Ser Glu Asp Tyr Ala His Leu Lys Glu Pro
227 145 150 155 160
229 tgg ctg aag ctc acc ggc ttc ccc tgc tcc tgg acc ttc ttc cac cac 528
230 Trp Leu Lys Leu Thr Gly Phe Pro Cys Ser Trp Thr Phe Phe His His
231 165 170 175
233 ctc cgg gag cag atc cgc agc gag ttc acc ctg cac gac cac ctt cgg 576
234 Leu Arg Glu Gln Ile Arg Ser Glu Phe Thr Leu His Asp His Leu Arg
235 180 185 190
237 caa gag gcc cag ggg gta ctg agt cag ttc cgt cta ccc cgc aca ggg 624
238 Gln Glu Ala Gln Gly Val Leu Ser Gln Phe Arg Leu Pro Arg Thr Gly
239 195 200 205
241 gac cgc ccc agc acc ttc gtg ggg gtc cac gtg cgc cgc ggg gac tat 672
242 Asp Arg Pro Ser Thr Phe Val Gly Val His Val Arg Arg Gly Asp Tyr
243 210 215 220
245 ctg cgt gtg atg ccc aag cgc tgg aag ggg gtg gtg ggt gac ggc gct 720
246 Leu Arg Val Met Pro Lys Arg Trp Lys Gly Val Val Gly Asp Gly Ala

```

## RAW SEQUENCE LISTING

DATE: 07/07/2003

PATENT APPLICATION: US/09/051,034B

TIME: 14:36:20

Input Set : A:\305626u1.app

Output Set: N:\CRF4\07072003\I051034B.raw

```

247 225          230          235          240
249 tac ctc cag cag gct atg gac tgg ttc cgg gcc cga tac gaa gcc ccc 768
250 Tyr Leu Gln Gln Ala Met Asp Trp Phe Arg Ala Arg Tyr Glu Ala Pro
251          245          250          255
253 gtc ttt gtg gtc acc agc aac ggc atg gag tgg tgc cgg aag aac atc 816
254 Val Phe Val Val Thr Ser Asn Gly Met Glu Trp Cys Arg Lys Asn Ile
255          260          265          270
257 gac acc tcc cgg ggg gac gtg atc ttt gct ggc gat ggg cgg gag gcc 864
258 Asp Thr Ser Arg Gly Asp Val Ile Phe Ala Gly Asp Gly Arg Glu Ala
259          275          280          285
261 gcg ccc gcc agg gac ttt gcg ctg ctg gtg cag tgc aac cac acc atc 912
262 Ala Pro Ala Arg Asp Phe Ala Leu Leu Val Gln Cys Asn His Thr Ile
263          290          295          300
265 atg acc att ggc acc ttc ggc ttc tgg gcc gcc tac ctg gct ggt gga 960
266 Met Thr Ile Gly Thr Phe Gly Phe Trp Ala Ala Tyr Leu Ala Gly Gly
267 305          310          315          320
269 gat acc atc tac ttg gct aac ttc acc ctg ccc act tcc agc ttc ctg 1008
270 Asp Thr Ile Tyr Leu Ala Asn Phe Thr Leu Pro Thr Ser Ser Phe Leu
271          325          330          335
273 aag atc ttt aaa ccc gag gct gcc ttc ctg ccc gag tgg gtg ggc att 1056
274 Lys Ile Phe Lys Pro Glu Ala Ala Phe Leu Pro Glu Trp Val Gly Ile
275          340          345          350
277 aat gca gac ttg tct cca ctc cag atg ttg gct ggg cct tga 1098
278 Asn Ala Asp Leu Ser Pro Leu Gln Met Leu Ala Gly Pro
279          355          360          365
282 <210> SEQ ID NO: 4
283 <211> LENGTH: 365
284 <212> TYPE: PRT
285 <213> ORGANISM: Sus Domesticus
287 <400> SEQUENCE: 4
288 Met Trp Val Pro Ser Arg Arg His Leu Cys Leu Thr Phe Leu Leu Val
289 1 5 10 15
290 Cys Val Leu Ala Ala Ile Phe Phe Leu Asn Val Tyr Gln Asp Leu Phe
291 20 25 30
292 Tyr Ser Gly Leu Asp Leu Leu Ala Leu Cys Pro Asp His Asn Val Val
293 35 40 45
294 Ser Ser Pro Val Ala Ile Phe Cys Leu Ala Gly Thr Pro Val His Pro
295 50 55 60
296 Asn Ala Ser Asp Ser Cys Pro Lys His Pro Ala Ser Phe Ser Gly Thr
297 65 70 75 80
298 Trp Thr Ile Tyr Pro Asp Gly Arg Phe Gly Asn Gln Met Gly Gln Tyr
299 85 90 95
300 Ala Thr Leu Leu Ala Leu Ala Gln Leu Asn Gly Arg Gln Ala Phe Ile
301 100 105 110
302 Gln Pro Ala Met His Ala Val Leu Ala Pro Val Phe Arg Ile Thr Leu
303 115 120 125
304 Pro Val Leu Ala Pro Glu Val Asp Arg His Ala Pro Trp Arg Glu Leu
305 130 135 140
306 Glu Leu His Asp Trp Met Ser Glu Asp Tyr Ala His Leu Lys Glu Pro

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 07/07/2003  
PATENT APPLICATION: US/09/051,034B      TIME: 14:36:21

Input Set : A:\305626u1.app  
Output Set: N:\CRF4\07072003\I051034B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:16; Xaa Pos. 3,4